

# 第三部分 HDR 编码方案

## 第十章-MPEG 及 VCEG 的 HDR 编码优化

与传统标准动态范围（SDR）视频相比，高动态范围（HDR）视频由于比特深度的增加提供了更加丰富的亮区细节和暗区细节。最新的显示技术通过清晰地再现 HDR 视频内容使得为用户提供身临其境的观看体验成为可能。面对目前日益增长的 HDR 视频消费需求，研究现有的压缩工具或引入新的技术来高效压缩 HDR 视频变得十分迫切。本文将介绍有关 MPEG 及 VCEG 针对 HDR 视频压缩的一些工作。

### 10.1 背景

传统电视和显示器上显示的视频信号为标准动态范围视频，设计的亮度范围为  $0.01\text{cd/m}^2$  到  $100\text{cd/m}^2$ ，对比度为 10000: 1。而在现实世界中，亮度范围要大得多，例如阳光照射的云层具有  $100000\text{cd/m}^2$  的亮度，而夜空的亮度仅为  $0.0001\text{cd/m}^2$ 。近年来，显示器的亮度峰值持续在增加，市面上已经出现了亮度峰值可以达到  $1000\text{cd/m}^2$  的电视机产品，而一些专业的监视器则可以达到  $10000\text{cd/m}^2$ ，这些显示器大多基于 LCD 技术以及 LED 背光，使得亮度峰值能够达到较高的水平。OLED 显示器也可以达到一个很高的亮度级别，但峰值略低于 LCD & LED 技术，不过 OLED 的优点是其暗度可以达到  $0.0005\text{cd/m}^2$ 。

超高清电视（UHDTV）的关键目标之一是增强用户观看视频时的体验，也就是所谓的“沉浸感”和“真实感”。仅仅依靠提高分辨率不足以完全达到这一目的，也无法创建、捕捉和显示比今天的电视机具有更高峰值亮度和更大对比度的视频内容。高动态范围（HDR）视频的目的是提高每个像素还原视频真实场景内容的能力，特别是，最亮像素的亮度增加，同时最暗像素可以变得更暗。另外，呈现更强的真实感需要提供更加丰富的色彩，因此必须使用具有较大色域范围的颜色空间（宽色域 WCG）来表示更加饱满的色彩。因此，高动态范围（HDR）与宽色域（WCG）相结合，可以使视频看起来更加清晰、逼真，与传统的标准动态范围视频相比，能够有效改善用户的视频观看体验。

与 SDR 视频相比，可以从以下几个方面了解 HDR 视频的效果：

1. 即使分辨率没有改变，图像也可能看起来更加清晰。人眼将清晰度视为分辨率和对比度的组合。因此，当对比度增加时，人眼感知的清晰度也会增加。因此，HD 和 4K 视频都可以通过高动态范围（HDR）增强视觉效果。
2. 明亮的物体看起来会更加逼真。许多物体包含明亮区域，即使很小，也可能会影响观看效果，因为它们比周围环境更亮。HDR 视频会以较高的保真度地表现明亮区域的细节，增加水面、玻璃表面、金属表面等物体的真实感。
3. 阴影区域中的细节可能会增加。在 SDR 视频中，较暗区域看起来是均匀和平坦的，而在 HDR 视频相同场景下可以清楚地看到阴影中的细节。
4. HDR 和 WCG 的组合还会产生更加饱满的色彩。对于诸如天空之类的明亮物体，HDR 和 WCG 允许出现明亮且同时饱满的颜色。

## 10.2 HDR 发展历程

### 10.2.1 ITU-R for SDR and HDR

在 ITU-R 的建议书中描述了用于表示 SDR 和 HDR 视频格式的基本技术规范。其中 BT.709 “Parameter values for the HDTV standards for production and international programme exchange” 包括了如何将 8 比特线性的 RGB 图像信号转换为非线性的 R'G'B' 信号，这个转换通常称为光电转换函数 (opto-electronic transfer function, OETF)。BT.709 还包含了 CIE1931 色度坐标的定义，以及从 R'G'B' 向 Y'CbCr 颜色空间转换方法的介绍。实际上，相比于色度的变化，人眼对于亮度的变化更为敏感，因此通过这种转换之后就可以降低色度分量的空间采样率。但应该注意的是，亮度分量并不直接对应于实际的亮度信息，一部分亮度信息包含在色度分量中，这就会导致在对色度分量进行下采样时也会对亮度信息产生影响，这种现象称为色度泄漏 (chroma leakage)，这对 HDR 视频的影响会比 SDR 视频更严重，下文会讨论关于色度泄漏的解决方法。BT.1886 “Reference electro-optical transfer function for flat panel display used in HDTV studio production” 中介绍了从非线性 R'G'B' 信号向线性 RGB 信号转换的操作方法，该转换操作在视频内容显示前进行，称为电光转换函数 (electro-optical transfer function, EOTF)。这里提出的电光转换方程 EOTF 与 BT.709 标准中的光电转换方程 OETF 并不是互逆的关系，此 EOTF 和逆 OETF 转换函数的差异可以用一个光电光学转换函数 (opto-optical transfer function) 或系统参数 gamma 来表示，该 gamma 参数能够使得图像看起来更加自然真切。BT.2035 “A reference viewing environment for evaluation of HDTV programme material or completed programmes” 中给出了视频观看环境的参考设置，其中规定了亮度范围为  $0.01\text{cd/m}^2$  到  $100\text{cd/m}^2$ 。

BT.2020 “Parameter values for ultra-high definition television systems for production and international programme exchange” 中引入了 4K (3840x2160) 分辨率和宽色域 (WCG) 的概念。宽色域意味着其基色范围相对于传统视频更广，因而可以显示出更加饱满的颜色，例如深蓝、品红、荧光绿等。但需要注意的是，没有 HDR 的 WCG 其效果是很有限的，同样地，不引入 WCG 的 HDR 视频也是没有太大意义的，因此 HDR 和 WCG 可以很自然地结合起来用于改善每个像素点的重建效果，呈现出更为生动的视频内容。此外，如前文所述，HDR 和 WCG 产生的效果几乎完全独立于空间分辨率，因此可以灵活应用在 4K 分辨率及 1080p 分辨率视频中。与此同时，BT.2020 包括了与 BT.709 相同的非恒定亮度颜色空间转换，还包括可以直接应用在 RGB 信号上的恒定亮度转换。在恒定亮度下，OETF 在使用前要先进行颜色空间转换，但这种方法没有非恒定亮度那样应用广泛。

BT.2100 “Image parameter values for high dynamic range television for use in production and international programme exchange” 中介绍了两种不同的 HDR 转换函数：HLG (Hybrid Log-Gamma) 和 PQ (perceptual quantization)。HLG 的光电转换函数 OETF 与 ARIB STD-B67 中的 OETF 是相同的，并且对支持 WCG (BT.2020) 的 SDR 显示器有一定的兼容性，但需要注意的是 HLG 本身不兼容只支持 BT.709 色域的显示器。PQ 电光转换函数 EOTF 与 SMPTE ST.2084 中的 EOTF 相同，并具有与人类视觉系统中分布比较一致的亮度级。SMPTE ST.2084 于 2013 年发布，近年来被广泛应用于工业界，并被多个不同的标准组织采纳。MPEG 和 JCT-VC 相关的工作已经将 PQ 电光转换函数 EOTF 作为研究和测试的基准。此外，BT.2100 中的颜色基色与颜色空间转换和 BT.2020 中给出的一致。

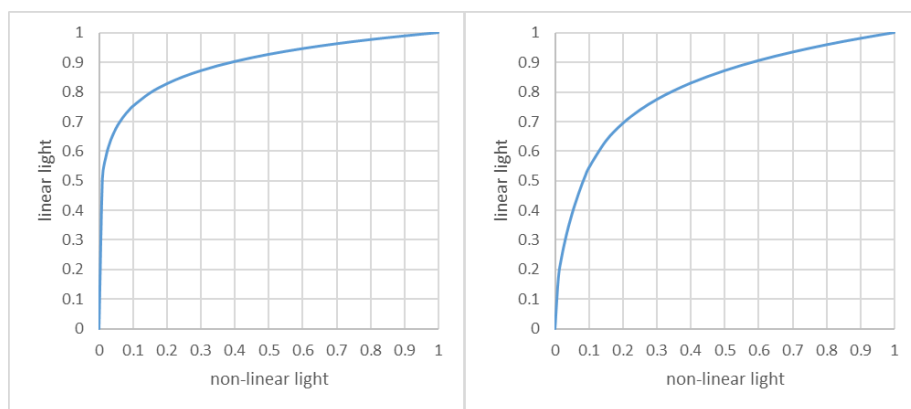


图 10-1 PQ（左）和 HLG（右）转换曲线

## 10.2.2 MPEG 和 VCEG 关于 HDR 的工作

MPEG 在 2013 年开始研究 HDR 视频压缩时进行了多项探索性的研究测试，主要思路是将 HDR 视频压缩和处理过程的各部分分开来研究，例如在此期间研究了不同的电光转换函数，并对不同的颜色格式转换和采样方法进行了测试和评估，与此同时还研究了一些主观视频质量测试方法和客观视频质量评价指标。可以预计的是，传统的基于像素误差的评价指标如 PSNR 在高精度 HDR 视频的评价中无法提供太多有价值的参考信息。目前已提出了一些专门针对 HDR 视频的质量指标，例如 wPSNR、tPSNR-Y、deltaE100、PSNR-L100 等，但仍然没有一个主客观评估结果较为一致的评价指标。

2015 年初，MPEG 决定发布针对 HDR 和 WCG 视频压缩技术标准的 CfE[1]，目的是为了确定当前是否存在能够改善 HDR 视频压缩性能的新技术。CfE 中给出了一个参考配置作为对比基准 Anchor 1.0，其中大部分技术方案与 HDR10（在下一节介绍）完全兼容。Anchor 1.0 使用 PQ 电光转换函数、Y'CbCr 颜色空间、HEVC Main10 压缩配置，基于未修改的 HEVC 参考软件 HM 进行编码，并使用 HDR Tools 软件包[2]来实现必要的预处理和后处理转换步骤。HDR Tools 本身是开源免费的，其在 MPEG 和 VCEG 进行的多种评估测试中都有所应用。

CfE 提出后，收到了来自 Apple、Arris、BBC、Dolby、Ericsson、FastVDO、InterDigital、MovieLabs、NGCodec、Philips、Qualcomm、Technicolor、University of Warwick/goHDR 等总共 9 种不同的技术提案，所有提案的编码结果在两个独立的实验室分别进行了主观评测，得到的结果与 Anchor 1.0 进行了对比。主观评测结果显示，在很多序列上部分提案的结果相比 Anchor 1.0 有明显的提升[3]。

与此同时，CfE 中的一些序列测试结果表明，可以通过 chroma QP offset [4]（第 3 节介绍）将少量比特从亮度分量转移到色度分量，从而可以减少色度失真。此外也发现在 Anchor 1.0 中分配给图像暗区的比特过多，而分配给亮区的比特太少[5]，为此引入了亮度调整（Luma Adjustment，第 3 节介绍）方法。最后综合多种新的改进技术形成了另一个基准，Anchor 3.2。相对于 Anchor 1.0，Anchor 3.2 在各个方面都有很大程度的改善。与 Anchor 1.0 中的方法相比，同等视觉质量条件下 Anchor 3.2 可以实现大约平均 27% 的比特率降低（使用 MOS BD-Rate 计算）。不过应该注意的是，Anchor 3.2 中的改变都发生在 HEVC 编码端，而在解码端没有作任何改变，也就是说 Anchor 3.2 与 HDR10 完全兼容。Anchor 3.2 与 Anchor 1.0 的编码结果对比如图 2 所示。



图 10-2 Anchor 1.0（左）和 Anchor 3.2（右）在相同码率下的压缩效果

在 2016 年 2 月的 MPEG 会议上，对已有工作进行了分析后得出结论：暂时没有找到一套明确的技术方案能够依此建立针对 HDR 视频的新的 HEVC 规范，因此研究重点转向了技术报告。MPEG 和 VCEG 随后的研究工作明确了与 HDR 视频转换和压缩相关的技术特性，并形成了一份记录这些特性的技术报告，用来指导 HDR 视频的处理过程，即对支持 PQ 转换函数、4:2:0 Y'CbCr 并使用 HEVC Main10 Profile 的 HDR 编码方案，也就是 Anchor 3.2 中所包含的编码 HDR 视频的方案提供技术参考。该技术报告从 2015 年就开始着手准备了，第一个正式的版本在 2016 年 2 月发布，随后被确定为 MPEG-H 的一个新的部分，也就是 MPEG-H Part 14。该技术报告的第一版为“Conversion and coding practices for HDR/WCG video” [6]，后来标题变得更加具体，即为“Conversion and Coding Practices for HDR/WCG Y'CbCr 4:2:0 Video with PQ Transfer Characteristics” [7]。

这份技术报告所描述的 HDR 视频处理流程如图 3 所示，其中包括 BT.2020 中所规定的线性 RGB 图像信号向 10 比特、窄范围、PQ 转换函数、4:2:0 采样率、非恒定亮度 Y'CbCr 图像信号进行转换时所必需的预处理步骤，具体包括 EOTF 的应用、颜色空间转换、量化和色度下采样等转换操作。相应地，该报告还介绍了在 HDR 视频解码之后执行的后处理转换步骤，这些步骤用于将视频信号转换回线性的 RGB 表示，从而能够在显示器上显示。此外，中间编码部分的介绍并不是对视频编码的完整描述，只是突出说明了针对 HDR 视频压缩的两种有效的方法：基于平均亮度水平的 QP 自适应选择和基于内容色域的色度 QP 偏移，这两种方法将在本文第三部分详细介绍。

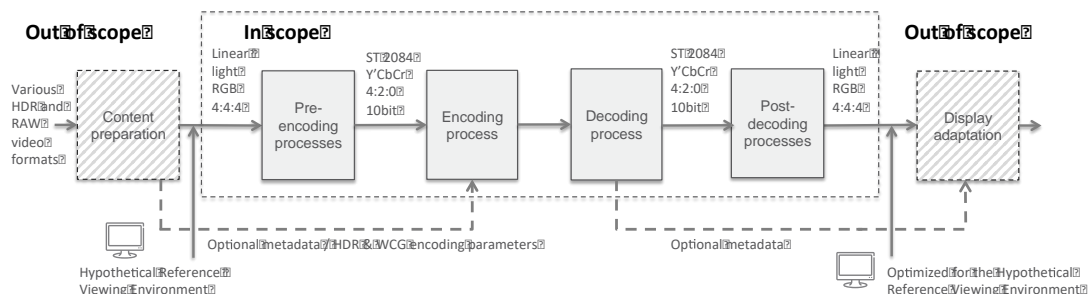


图 10-3 HDR 视频处理流程<sup>[7]</sup>

## 10.2.3 基于 HEVC 的 HDR 方案

HEVC 具有较高的压缩效率，支持 10 比特精度、HDR 转换函数和 WCG，是 HDR/WCG 视频数据的理想载体。[8]中给出了将 HEVC 编码器及其 Main10 配置文件用于消费者 HDR/WCG 应用的可行性。该报告记录了使用 HEVC Main10 配置文件和 BT.2100 PQ 转换函数在不同比特率下编码 HDR 后所进行的一些主观测试结果。该报告的结论是，HEVC Main 10 配置文件确实能够以行业常见的比特率高质量地编码 HDR/WCG 视频，而无需对 HEVC 标准进行任何进一步的修改。[7]中具体给出了关于如何使用 HEVC 和 PQ 函数转换 HDR 视频以进行高效压缩，而[9]中给出了一些使用 HEVC 和 HLG 函数进行 HDR 视频编码的指导性方案。

HEVC 编码器、10 比特精度、BT.2100 色彩基色和 PQ 转换函数的组合通常称为 PQ10，同样，HEVC 编码器、10 比特精度、BT.2100 色彩基色和 HLG 转换函数通常称为 HLG10。PQ10 和 HLG10 编码 HDR 的方案目前已被许多标准化机构采用和部署，其中包括：

- ETSI / DVB

2017 年 2 月发布的技术规范 2.3.1 版本中，UHD 第 2 阶段部署采用 HEVC 编码，包括使用 PQ10 或 HLG10 方案。

- BD ROM

BDA (Blu-ray Disc Association) 是最早支持 HDR 视频的行业组织之一，编码方案中采用了多个不同的 HDR 关键技术，其中主要使用的是 HDR10。HDR10 技术涵盖：PQ 转换函数、BT.2020 色域、非恒定亮度 Y'CbCr 颜色空间、4:2:0 采样格式、采用 HEVC Main10 编码配置。在实际应用时，HDR10 中还包括一些常用的元数据，比如 Mastering Display Colour Information 及 Content Light Level Information 等。

2015 年 7 月发布的用于蓝光光盘的技术规范 (3.0 版本) 支持 HEVC Main 10 Profile Level 5.1，最高码率 100 Mbps，分辨率为 HD 或 4K，帧率 24-60fps，兼容 PQ10 方案。

- Ultra HD Forum

2017 年 4 月 1 日发布的技术规范“UHD Phase A 内容参数”指定使用 HEVC Main10 编码配置，并采用 PQ10 或 HLG10 中的颜色表示方案。

- 3GPP

基于感知量化器 PQ 并遵循 ITU-R BT.2100 颜色格式的 HDR 方案被添加到了 3GPP TS 26.116 电视视频配置集中，具体包括 HEVC Main10 编码方案，转换函数、颜色基色和相关元数据的规定，以及与 PSS (分组交换流) 和 MBMS (多媒体广播和多播服务) 的集成。

- ATSC 3.0

使用 PQ10 或 HLG10 方案编码 HDR 视频。

除了标准化机构采用 PQ10 和 HLG10 之外，还有许多商业机构直接使用 HEVC 标准进行 HDR/WCG 编码、传输和播放，其中包括行业领先的流媒体服务商、主流的手机和电视制造商。目前基本上市场上所有销售的具有 HDR 功能的电视机都内置了 10 比特 HEVC 解码器。



## 10.3 HDR 视频编码端的关键技术

### 10.3.1 Luma adjustment

之前已经提到，在  $Y'CbCr$  表示中，一些亮度信息携带在色度分量中，当色度分量进行下采样时，可能出现色度泄漏。对于 HDR 视频，由于 PQ 电光转换函数 EOTF 具有高度非线性特性以及相邻像素之间的亮度可能显著不同，这种影响可能非常大。图 4 给出了色度泄漏的示例，并在之后介绍了如何使用亮度调整方法解决此问题。应当注意，该例子使用了 BT.709 容器中所有可用的色度代码值，并在 BT.709 显示器上分级显示。这是因为在 MPEG 中暂时没有可用的具有完整 BT.2020 色域的测试材料，但是对于 BT.2020 容器中的测试材料预期也会具有类似的效果。因此，图 4 中的示例经过了色调映射，以便能够在 SDR 显示器中呈现。



图 10-4 色度泄漏示例 (a.原图 (4:4:4) b.色度下采样图 (4:2:0) c. Luma adjustment)

如图 5 所示，亮度调整方法[10]增加了色度上采样环节，使得  $Cb$  和  $Cr$  信息参与到亮度分量的调整中来，以确保解码后在对色度进行上采样时保留有价值的亮度信息，如图 6 所示。

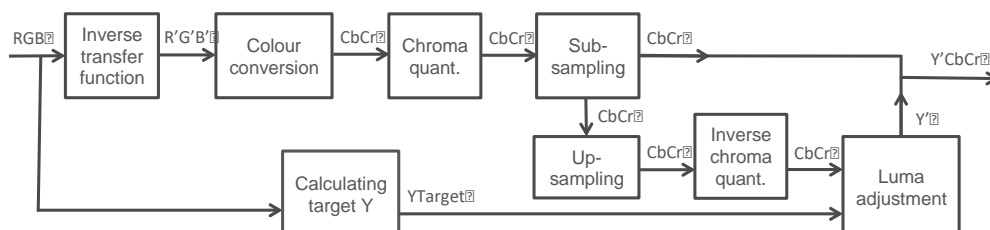


图 10-5 编码预处理过程亮度调整方法示意图

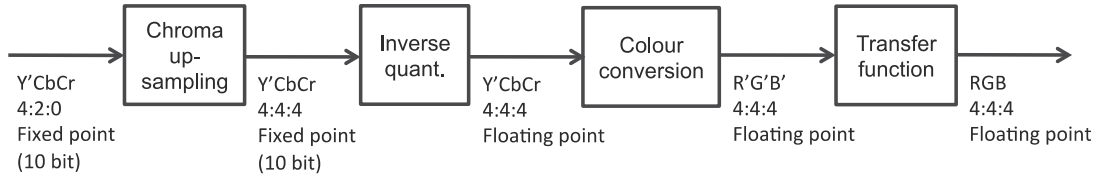


图 10-6 解码后处理过程示意图

在后处理阶段，根据亮度调整值 $Y'_{\text{adjust}}$ 和经过上采样反量化得到的 $\widetilde{\text{Cb}}$ 和 $\widetilde{\text{Cr}}$ 经过一系列转换可以得到重建值 $\{\widetilde{\text{R}}, \widetilde{\text{G}}, \widetilde{\text{B}}\}$ ，确保 $\widetilde{\text{RGB}}$ 能够接近最初的RGB是优化的目标。 $\widetilde{\text{RGB}}$ 和RGB之间的差异可以表示如下：

$$D = \text{Diff}(\text{RGB}, \widetilde{\text{RGB}}) \quad (1)$$

一般采用基于权重的误差和来表示D，权重 $w_R, w_G, w_B$ 可以设定为各分量对亮度的贡献比例。

$$D = \left( w_R * (\widetilde{\text{R}} - \text{R}) + w_G * (\widetilde{\text{G}} - \text{G}) + w_B * (\widetilde{\text{B}} - \text{B}) \right)^2 \quad (2)$$

根据R和R'之间的EOTF转换关系，上式等价于：

$$D = \left( w_R * \left( \text{EOTF}(\widetilde{\text{R}}') - \text{EOTF}(\text{R}') \right) + w_G * \left( \text{EOTF}(\widetilde{\text{G}}') - \text{EOTF}(\text{G}') \right) + w_B * \left( \text{EOTF}(\widetilde{\text{B}}') - \text{EOTF}(\text{B}') \right) \right)^2 \quad (3)$$

其中， $\text{EOTF}(x)$ 为PQ EOTF。为了得到可以收敛的结果，对 $\text{EOTF}(x)$ 进行泰勒级数展开：

$$\text{EOTF}(x_i + \Delta) \approx \text{EOTF}(x_i) + \text{EOTF}'(x_i) * \Delta_x \quad (4)$$

$\text{EOTF}'(x_i)$ 是 $\text{EOTF}(x)$ 关于x在 $x_i$ 处的一阶导数。将(4)式带入(3)式可得：

$$D = (w_R * \text{EOTF}'(\text{R}') * \Delta_R + w_G * \text{EOTF}'(\text{G}') * \Delta_G + w_B * \text{EOTF}'(\text{B}') * \Delta_B)^2 \quad (5)$$

由于 $\text{R}'\text{G}'\text{B}'$ 与 $Y'\widetilde{\text{Cb}}\widetilde{\text{Cr}}$ 之间可以进行如下转换：

$$\begin{bmatrix} \text{R}' \\ \text{G}' \\ \text{B}' \end{bmatrix} = \mathbf{A} \begin{bmatrix} Y' \\ \text{Cb} \\ \text{Cr} \end{bmatrix} = \mathbf{W}^{-1} \begin{bmatrix} Y' \\ \text{Cb} \\ \text{Cr} \end{bmatrix} = \frac{1}{\det(\mathbf{W})} \text{adj}(\mathbf{W}) \begin{bmatrix} Y' \\ \text{Cb} \\ \text{Cr} \end{bmatrix} = \begin{bmatrix} a_{RY} & a_{RCb} & a_{RCr} \\ a_{GY} & a_{Gcb} & a_{GCr} \\ a_{BY} & a_{BCb} & a_{BCr} \end{bmatrix} \begin{bmatrix} Y' \\ \text{Cb} \\ \text{Cr} \end{bmatrix} \quad (6)$$

将上式代入(5)式替换掉变化量，D可表示为如下所示，记为(7)式。

$$\begin{aligned} & (w_R * \text{EOTF}'(\text{R}') * (a_{RY} * Y'_{\text{adjust}} + a_{RCb} * \widetilde{\text{Cb}} + a_{RCr} * \widetilde{\text{Cr}} - (a_{RY} * Y' + a_{RCb} * \text{Cb} + a_{RCr} * \text{Cr})) + \\ & w_G * \text{EOTF}'(\text{G}') * (a_{GY} * Y'_{\text{adjust}} + a_{Gcb} * \widetilde{\text{Cb}} + a_{GCr} * \widetilde{\text{Cr}} - (a_{GY} * Y' + a_{Gcb} * \text{Cb} + a_{GCr} * \text{Cr})) + \\ & w_B * \text{EOTF}'(\text{B}') * (a_{BY} * Y'_{\text{adjust}} + a_{BCb} * \widetilde{\text{Cb}} + a_{BCr} * \widetilde{\text{Cr}} - (a_{BY} * Y' + a_{BCb} * \text{Cb} + a_{BCr} * \text{Cr})))^2 \end{aligned}$$

经过矩阵A的系数替换之后，可简化为：

$$D = \left( w_R * \text{EOTF}'(\text{R}') * (Y'_{\text{adjust}} - e_R) + w_G * \text{EOTF}'(\text{G}') * (Y'_{\text{adjust}} - e_G) + w_B * \text{EOTF}'(\text{B}') * (Y'_{\text{adjust}} - e_B) \right)^2 \quad (8)$$

其中部分系数如下所示：

$$e_R = Y' - (\widetilde{Cr} - Cr) * a_{RCr}, \quad (9)$$

$$e_G = Y' - (\widetilde{Cb} - Cb) * a_{GCB} - (\widetilde{Cr} - Cr) * a_{GCr}, \quad (10)$$

$$e_B = Y' - (\widetilde{Cb} - Cb) * a_{BCb}. \quad (11)$$

为了得到(8)式的局部最优解，对其关于 $Y'_{\text{adjust}}$ 求导并令D为0可得：

$$Y'_{\text{adjust}} = \frac{w_R * \text{EOTF}'(R') * e_R + w_G * \text{EOTF}'(G') * e_G + w_B * \text{EOTF}'(B') * e_B}{w_R * \text{EOTF}'(R') + w_G * \text{EOTF}'(G') + w_B * \text{EOTF}'(B')} \quad (12)$$

计算(12)式即可求导调整后的亮度值 $Y'_{\text{adjust}}$ 。

### 10.3.2 QP selection based on average luma

对于 SDR 数据，使用的是 BT.709 中的幂律转换函数，而 HDR/WCG 数据则一般使用在 SMPTE ST 2084 和 BT.2100 中定义的 PQ 转换函数，因此经过处理后数据特性会有很大差别，可以使用 BT.709 转换函数和 BT.2100 PQ 转换函数处理相同的 SDR 数据来观察这一差异。对于 10 比特的数据，如果源数据的亮度峰值为 100cd/m<sup>2</sup> 并使用 BT.709 转换函数，亮度分量分布在 64-940，而如果使用 BT.2100 PQ 转换函数，亮度分量分布在 64-590，可以看到两者范围差异较大。如果使用 PQ 时在 code level 为 509 (100cd/m<sup>2</sup>) 附近出现一些扰动 (+/- 1)，等价于使用 BT.709 在 code level 为 940 (100cd/m<sup>2</sup>) 出现 +/- 4 的扰动。与此同时，使用 PQ 时在 code level 为 80 (0.01cd/m<sup>2</sup>) 附近出现一些扰动 (+/- 1)，等价于使用 BT.709 在 code level 为 80 (0.01cd/m<sup>2</sup>) 出现 +/- 1 的扰动。因此，可以推断如果编码器在亮区和暗区附近使用相同的方式处理 1 个 code level 的误差，那么相对 BT.709 使用 PQ 时将会在亮区 (100cd/m<sup>2</sup>) 产生 4 倍左右的误差。也就是说，从 BT.709 切换到 PQ 将会导致许多比特从亮区重新分配到暗区。因此，原先针对 BT.709 在亮区和暗区取得较好平衡的编码器，如果不做修改而直接使用 PQ，则分配给亮区的比特过少，暗区的比特过多，所以亮区就会出现失真，对于 HDR/WCG 数据会表现得更加明显。一种常见的做法是基于局部块的平均亮度值自适应地调整块的 QP。

许多编码器都采用了自适应 QP 选择的方法。在 HEVC 中，可以单独为每个编码单元设置 QP，以便使得比特分配与人类视觉系统特性保持一致。通常，对于平坦区域（低方差）使用较低的 QP 并且对于纹理密集的区域（高方差）使用较高的 QP。此外在其他一些系统中，还可能会考虑亮度、边缘、运动等特性。但这些方法都是针对 SDR 内容特性的。对于使用 PQ 电光转换函数的 HDR 视频，需要考虑上述转换函数带来的数据特性上的改变。具体地，通过对具有较低平均亮度值的块（暗区域）使用较高的 QP，而对具有较高平均亮度值的块（亮区域）使用较低的 QP，可以有效改善图像的主观质量。下面给出一个基于平均亮度进行自适应量化的示例。

$$QP_{PQ} = QP_{709} + dQP(int_L) \quad (13)$$

$int_L$  表示局部 64x64 CTU 编码块的平均亮度值  $L_{\text{average}}$ ， $int_L = \text{Round}(L_{\text{average}})$ 。dQP 与  $int_L$  的对应关系通过如下查找表和图 7 给出。

表 1 dQP 与  $int_L$  之间的对应关系

luma $int_L$ range	dQP
$int_L < 301$	3
$301 \leq int_L < 367$	2
$367 \leq int_L < 434$	1
$434 \leq int_L < 501$	0
$501 \leq int_L < 567$	-1



$567 \leq \text{int}_L < 634$	-2
$634 \leq \text{int}_L < 701$	-3
$701 \leq \text{int}_L < 767$	-4
$767 \leq \text{int}_L < 834$	-5
$\text{int}_L \geq 834$	-6

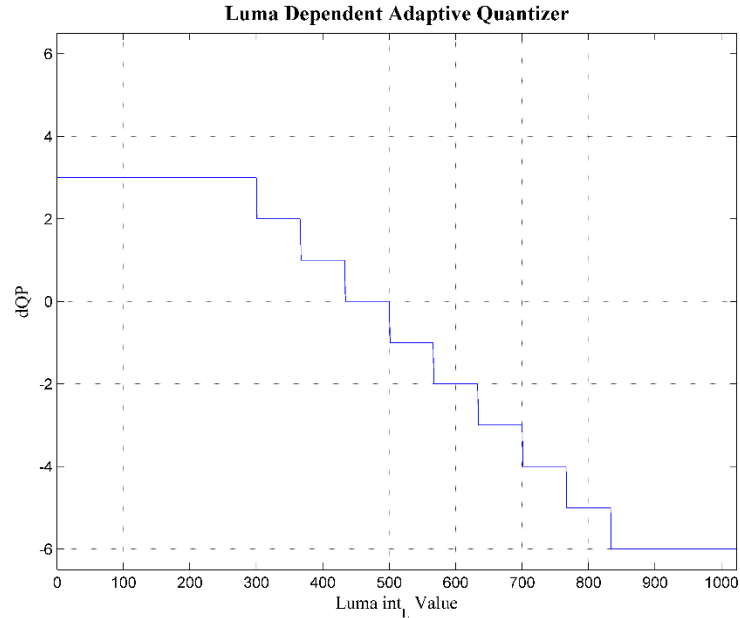


图 10-7 dQP 与 int<sub>L</sub> 之间的对应关系

### 10.3.3 Chroma QP offset based on content colour gamut

针对使用 BT.709 色域和光电转换函数的 SDR 内容，现有的编码器配置已经可以在亮度和色度之间取得较好的平衡。对于 HDR 视频，与 SDR 视频相比，PQ EOTF 和 BT.2020 色彩基色的组合会使得 Cb 和 Cr 的取值更加紧密地聚集在零附近。因此，使用同样的编码配置编码 HDR 视频可能不会获得与 SDR 类似的性能。具体在编码 HDR 视频时，过于集中的色度分量在量化过程中会出现丢失，此时比特分配会从色度分量向亮度分量偏移。如果没有为色度分量分配足够多的比特，就会产生色度失真，例如，白色区域中会出现青色和品红色，如图 2 左图所示，窗户和伞内区域发散出了多种颜色。

HEVC 中的 Chroma QP offset，调节着亮度 QP 与色度 QP 之间的平衡关系。通过对 Cb 和 Cr 单独设置色度 QP 偏移，可以确保在亮度使用较高的 QP 值时，色度 QP 仍然维持在一个正常的水平，使得在量化期间不会发生较为严重的色度信息丢失。例如，在失真较为严重的低码率情况，对 Chroma QP offset 取一个较大的负值能够有效减少色度失真，在码率增加到某个点时，色度分量的质量可能已经足够好了，所以在高码率情况下可以采取较小的偏移值，甚至设为 0。另外，如果在编码过程中已知视频内容没有应用全色域（通过分析视频或视频附带的元数据可知），则可以进一步优化色度 QP 偏移以提供更好的编码结果。

以下给出了设置 chroma QP offset 的一个例子。假定已知 HDR 捕捉和显示设备的颜色基色，chroma QP offset 可以基于亮度 QP 值和基色因子  $c_{cb}c_{cr}$  采用下式给出的模型进行计算：

$$\text{QPoffsetCb} = \text{Clip3}\left(-12, 0, \text{Round}(c_{cb} * (k * \text{QP} + l))\right) \quad (14)$$

$$QPoffsetCr = Clip3(-12, 0, Round(c_{cr} * (k * QP + l))) \quad (15)$$

如果 HDR 捕捉和显示设备的颜色基色相同，则  $c_{cb} = 1$ ， $c_{cr} = 1$ 。如果捕捉设备采用 P3D65 基色，显示设备采用 BT.2020，则  $c_{cb} = 1.04$ ， $c_{cr} = 1.39$ 。如果捕捉设备采用 BT.709，显示设备采用 BT.2020，则  $c_{cb} = 1.14$ ， $c_{cr} = 1.78$ 。最后， $k = -0.46$ ， $l = 0.26$ 。

#### 参考资料

- [1] A. Luthra, E. Francois, W. Husak, “Call for Evidence (CfE) for HDR and WCG Video Coding”, MPEG2014/N15083, 110th MPEG Meeting, Geneva, 2015.
- [2] A. Tourapis, HDRTools software package: <https://gitlab.com/standards/HDRTools/>
- [3] V. Baroncini, T. Ebrahimi, “Test Results of Call for Evidence (CfE) for HDR and WCG Video Coding”, N15350, June 2015.
- [4] J. Samuelsson, M. Pettersson, J. Ström, K. Andersson, Not Public: “Using chroma QP offset on HDR sequences”, MPEG m36581, June 2015
- [5] J. Zhao, S.-H. Kim, A. Segall, K. Misra, Not Public: “Performance investigation of high dynamic range and wide color gamut video coding techniques”, MPEG m37439, Oct 2015.
- [6] J. Samuelsson, “Conversion and coding practices for high-dynamic-range and wide-colour-gamut video”, Draft 1, MPEG N16063, February 2016.
- [7] J. Samuelsson, C. Fogg, A. Norkin, A. Segall, J. Ström, G. Sullivan, A. Tourapis, “Conversion and Coding Practices for HDR/WCG Y’CbCr 4:2:0 Video with PQ Transfer Characteristics”, Draft 2, JCTVC-X1017, June 2016.
- [8] V. Baroncini, K. Andersson, A. K. Ramasubramonian, G. J. Sullivan, “Verification Test Report for HDR/WCG Video Coding Using HEVC Main 10 Profile”, N16273, June 2016
- [9] ITU-T Supplement H.Supp18 and ISO/IEC TR 23008-18 “Signalling, backward compatibility and display adaptation for HDR/WCG video coding”.
- [10] J. Strom, J. Samuelsson and K. Dovstam, “Luma Adjustment for High Dynamic Range Video”, Proceedings of the IEEE Data Compression Conference (DCC), March 2016.